

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE (REV 10-2000)		ATTORNEY'S DOCKET NUMBER MTN-029US
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C.371		U.S. APPLICATION NO. (If known, see 37 CFR 1.5) <div style="font-size: 1.5em; font-weight: bold; text-align: center;">10/018638</div>
INTERNATIONAL APPLICATION PCT/US99/07185	INTERNATIONAL FILING DATE 31 March 1999 (31.03.99)	PRIORITY DATE CLAIMED 01 April 1998 (01.04.98)
TITLE OF INVENTION GROWTH DIFFERENTIATION FACTOR-9 REGULATORY SEQUENCES AND USES THEREFOR		
APPLICANT(S) FOR DO/EO/US Martin MATZUK, et al.		
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:		
<ol style="list-style-type: none"> 1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C.371. 2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371. 3. <input type="checkbox"/> This is an express request to promptly begin national examination procedures (35 U.S.C. 371(f)). 4. <input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31). 5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2)) <ol style="list-style-type: none"> a. <input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau). b. <input checked="" type="checkbox"/> has been communicated by the International Bureau. c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US). 6. <input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C 371(c)(2)). 7. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)) <ol style="list-style-type: none"> a. <input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau). b. <input type="checkbox"/> have been communicated by the International Bureau. c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. d. <input checked="" type="checkbox"/> have not been made and will not be made. 8. <input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)). 9. <input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)). (unexecuted) (4 Sheets); 10. <input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)). 		
Items 11. to 16. below concern document(s) or information included:		
<ol style="list-style-type: none"> 11. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98; 12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included 13. <input type="checkbox"/> A FIRST preliminary amendment; <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment. 14. <input type="checkbox"/> A substitute specification. 15. <input type="checkbox"/> A change of power of attorney and/or address letter. 16. <input checked="" type="checkbox"/> Other items or information: PCT International Published Application (WO 99/50406) (without International Search Report) (37 sheets); Corrected Cover Sheet of PCT International Published Application (WO 99/50406) (with International Search Report) (5 sheets); International Preliminary Examination Report (10 sheets); Petition to Revive (2 sheets in duplicate); Checks in the amount of \$917 (application filing fee) and \$640 (Petition to Revive) based on small entity status; Certificate of First Class Mailing (1 sheet); and Return Postcard. 		

U.S. APPLICATION NO. (if known, see 37 CFR 1.53) <div style="font-size: 1.5em; font-weight: bold;">10/018638</div>		INTERNATIONAL APPLICATION NO <div style="font-weight: bold;">PCT/US99/07185</div>		ATTORNEY'S DOCKET NO <div style="font-weight: bold;">MTN-029US</div>	
-----------------------------------------------------------------------------------------------------------------------	--	--------------------------------------------------------------------------------------	--	-------------------------------------------------------------------------	--

17. <input checked="" type="checkbox"/> The following fees are submitted: BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) .(a/o November 1, 2000): Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO, and International Search Report not prepared by the EPO or JPO.....\$1040 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO\$890 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.455(a)(2)) paid to USPTO\$740 International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4).....\$710 International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4).....\$100 <div style="text-align: center;">ENTER APPROPRIATE BASIC FEE AMOUNT =</div>	CALCULATIONS PTO USE ONLY																															
	\$890.00																															
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input checked="" type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	\$130.00																															
<table border="1" style="width:100%; border-collapse: collapse;"> <tr> <th style="width:20%;">CLAIMS</th> <th style="width:20%;">NUMBER FILED</th> <th style="width:20%;">NUMBER EXTRA</th> <th style="width:20%;">RATE</th> <th style="width:20%;"></th> <th style="width:20%;"></th> </tr> <tr> <td>Total claims</td> <td style="text-align: center;">31-20 =</td> <td style="text-align: center;">11</td> <td style="text-align: center;">X \$18.00</td> <td style="text-align: center;">\$198.00</td> <td></td> </tr> <tr> <td>Independent claims</td> <td style="text-align: center;">7-3 =</td> <td style="text-align: center;">4</td> <td style="text-align: center;">X \$84.00</td> <td style="text-align: center;">\$336.00</td> <td></td> </tr> <tr> <td colspan="3">MULTIPLE DEPENDENT CLAIM(S) (if applicable)</td> <td style="text-align: center;">+ 280.00</td> <td style="text-align: center;">\$280.00</td> <td></td> </tr> <tr> <td colspan="4" style="text-align: center;">TOTAL OF ABOVE CALCULATIONS =</td> <td style="text-align: center;">\$1634.00</td> <td></td> </tr> </table>	CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE			Total claims	31-20 =	11	X \$18.00	\$198.00		Independent claims	7-3 =	4	X \$84.00	\$336.00		MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ 280.00	\$280.00		TOTAL OF ABOVE CALCULATIONS =				\$1634.00			
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE																													
Total claims	31-20 =	11	X \$18.00	\$198.00																												
Independent claims	7-3 =	4	X \$84.00	\$336.00																												
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ 280.00	\$280.00																												
TOTAL OF ABOVE CALCULATIONS =				\$1634.00																												
<input checked="" type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.	\$																															
SUBTOTAL =	\$917.00																															
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).	\$																															
TOTAL NATIONAL FEE =	\$																															
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property	\$																															
TOTAL FEES ENCLOSED =	\$917.00																															
	Amount to be: refunded	\$																														
	charged																															

a. ☒ Check in the amount of \$ 917.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ in the amount of \$ _____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 12-0080. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Jane E. Remillard, Esq.
LAHIVE & COCKFIELD, LLP
28 State Street
Boston, Massachusetts 02109
United States of America
(617)227-7400
Date: 19 December 2001

SIGNATURE
Jane E. Remillard
 NAME
38,220
 REGISTRATION NUMBER

10013638-0171502

PTO/PCT Rec'd 15 JUL 2002

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

#5

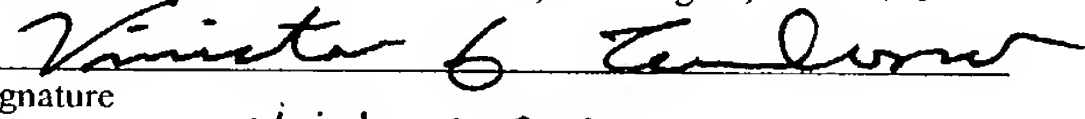
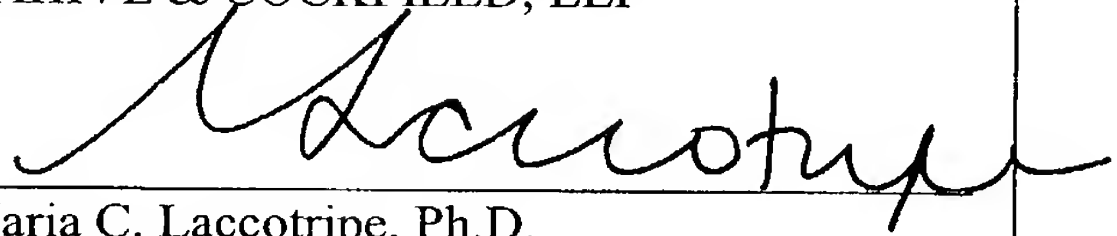
In re the application of: Martin MATZUK *et al.*
International Application No.: PCT/US99/07185
International Filing Date: 31 March 1999
U.S. Serial No.: 10/018,638
Filed: December 19, 2001
For: GROWTH DIFFERENTIATION FACTOR-9
REGULATORY SEQUENCES AND USES THEREFOR
Attorney Docket No.: MTN-029US

Commissioner for Patents
BOX SEQUENCE
Washington, D.C. 20231

TRANSMITTAL LETTER FOR DISKETTE CONTAINING SEQUENCE LISTING

Dear Sir:

Enclosed is a diskette which contains a computer readable form of the Sequence Listing for the above-referenced patent application. The Sequence Listing complies with the requirements of 37 C.F.R. § 1.821. The material on this diskette is identical in substance to the sequence listing appearing on pages 1- 3 of the Sequence Listing which is submitted herewith. The computer readable form of the Sequence Listing contained on the enclosed diskette is understood to comply with the requirements of § 1.824(d).

<p>Express Mail" mailing label number <u>EL 939 596 288 US</u></p> <p>Date of Deposit <u>July 15, 2002</u></p> <p>I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to: Box PCT, Commissioner for Patents, Washington, D.C. 20231</p> <p> Signature <u>Viriato G. Cardoso</u></p> <p>Please Print Name of Person Signing</p>	<p>Respectfully submitted,</p> <p>LAHIVE & COCKFIELD, LLP</p> <p> Maria C. Laccotripe, Ph.D. Limited Recognition Under 37 C.F.R. §10.9(b) Attorney for Applicants</p>
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

GROWTH DIFFERENTIATION FACTOR-9 REGULATORY SEQUENCES AND USES THEREFOR

Field of the Invention

5 The invention relates to tissue-specific regulatory elements derived from GDF-9 genes, as well as methods of identifying and using the regulatory elements to control gene expression in selected tissues.

Background of the Invention

10 Growth differentiation factor 9 (GDF-9) is a recently-identified member of the TGF β superfamily, which encompasses a large group of growth and differentiation factors that play important roles in regulating embryonic development and in maintaining tissue homeostasis in adult mammals. These factors are synthesized as pre-propeptides, forming dimers in the mature, secreted form. In this dimeric state, GDF
15 proteins are able to interact with various cell surface receptors, thereby modulating a variety of cellular processes. Several TGF β family members have been found to have tissue-specific expression related to the function of the factor. For example, GDF-8, a negative regulator of skeletal muscle growth, is specifically expressed in cells from skeletal tissue (McPherron, A.C., et al., Nature 387: 83-90, 1997).

20 The precise function of GDF-9 remains unclear, but the distribution of its gene product is known to be restricted to only a few tissues, particularly the ovary (and in the testes and the hypothalamus at very low levels). In humans, expression is also observed in the pituitary and placenta (Fitzpatrick et al., Endocrinology 139: 2571-2578, 1998). Studies in transgenic mice have found that disruption of the GDF-9 gene results in the
25 prevention of normal follicular development beyond an early stage, resulting in infertility in females (Dong et al., Nature 383: 531-535, 1996; Carabatsos et al., Developmental Biology 204: 373-384, 1998). No data is currently available about the role of GDF-9 in the testes, but male GDF-9-deficient mice exhibit normal fertility.

Accordingly, in another aspect, the present invention provides an expression vector containing one or more of the aforementioned isolated GDF-9 regulatory sequences, optionally linked to a gene of interest. The gene can be GDF-9 itself, or can be a gene whose product is desired to be under the control of one or more GDF-9 regulatory elements (e.g., a therapeutic gene or a reporter gene).

In another aspect, the invention provides a cell or tissue transfected to contain one or more of the aforementioned isolated polynucleotides. The cell can be, for example, a cell from the oocyte, testis, placenta, hypothalamus or pituitary.

In yet another aspect, the invention provides methods for obtaining or controlling (i.e., modulating) tissue-specific expression of a gene. The gene can be an endogenous gene or an exogenous gene (e.g., a transgene). In one embodiment, the method involves operatively linking one or more GDF-9 regulatory elements of the invention to a gene and then introducing (e.g., via microinjection) the resulting gene construct into a cell of a selected tissue, such as a tissue in which GDF-9 is naturally expressed (e.g., oocytes, testis, or hypothalamus). GDF-9 regulatory elements of the invention can be used both to upregulate gene expression in such tissues, and to downregulate (e.g., repress) gene expression in such tissues.

In yet another aspect, the invention provides a method for identifying tissue-specific regulatory elements derived from a GDF-9 gene. This can be achieved by, for example, comparing expression levels of a gene (e.g., a reporter gene), either *in vitro* or *in vivo* (e.g., in a transgenic animal), using different portions of the GDF-9 regulatory region spanning the 10 kilobase region immediately upstream (i.e., 5') of the coding region of a GDF-9 gene.

Additional aspects of the invention shall be apparent from the detailed
20 description below.

Brief Description of the Drawings

Figure 1 shows a comparison between the nucleotide sequence of the first 376 base pairs of the region immediately 5' of the transcription initiation site of the mouse (mGDF) and human (hGDF) GDF-9 genes. The sequences are oriented 3' to 5' so that the numbering of nucleotides (+1) begins after the ATG translation start codon. As shown, the overall identity between the mGDF and hGDF sequences is 71.3%.

Figure 2 shows a map of two GDF-9 transgenes constructed and utilized for site-specific expression studies. In panel A, the GFP transgene includes the genomic 10 kilobase 5'-flanking region immediately upstream of the start codon for GDF-9 operatively linked to the green fluorescence protein (eGFP) gene. In panel B, the GFP transgene includes the genomic 3.3 kilobase 5'-flanking region immediately upstream of

the start codon for GDF-9 operatively linked to the enhanced green fluorescence protein (eGFP) gene. The eGFP gene is a version of the GFP gene that has been optimized for mammalian expression and contains an optimized translation start site.

Figure 3 shows a northern analysis of tissues from a transgenic mouse line containing the 3.3 kilobase GDF-9-GFP construct shown in Figure 2. Tissues were hybridized with a ³⁵S-labeled anti-GFP probe. Both the ovary and the testis show marked GFP mRNA accumulation.

Figure 4 shows the results of a northern analysis of tissues from a different transgenic mouse line (compared to Figure 3) containing the 3.3 kilobase GDF-9-GFP construct shown in Figure 2. Tissues were hybridized with a ³⁵S-labeled anti-GFP probe. Both the ovary and the testis show marked GFP mRNA accumulation.

Figure 5 shows a northern analysis of tissues from a transgenic mouse line containing the 10 kilobase GDF-9-GFP construct shown in Figure 2. Tissues were hybridized with an anti-GFP probe. Only the ovary shows marked GFP mRNA accumulation.

Figure 6 shows a map comparing the 5' flanking sequences of the mouse GDF-9 gene and the mouse ZP3 gene (Lira et al. (1993) *Mol. Repro. Dev.* 36:494-499). Conserved regions, such as the E-Box at bases -182 and -183, respectively, are shown.

Figure 7 shows the sequencing strategy for the first 3.3 kilobases of the 5'-flanking sequence of the mouse GDF-9 gene. The numbering (+1) is relative to the transcription start site (ATG).

Detailed Description of the Invention

The present invention is based on the identification and isolation of GDF-9 regulatory elements which can be used to control (e.g., upregulate or down-regulate) expression of selected genes in specific tissues (e.g., oocyte, testis and other tissues in which GDF-9 is naturally expressed).

Definitions

As used herein, the terms defined below shall have the following meanings.

An "isolated polynucleotide" refers to a polynucleotide (e.g., a DNA) removed from its natural sequence context. The isolated polynucleotide can be any

polynucleotide that is capable of being transcribed or translated in a cell. The isolated polynucleotide can be, for example, cloned (genomic or cDNA clone) into a vector. Except as noted hereinafter, standard techniques for cloning, DNA isolation, amplification and purification, for enzymatic reactions involving DNA ligase, polymerase, restriction endonucleases and the like and various separation techniques are those known and commonly employed by those skilled in the art. A number of standard techniques are described in: Maniatis et al. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; Wu (ed.)(1979) Meth. Enzymol 68; Wu et al. (Eds.) (1983) Meth. Enzymol. 100 & 101; Grossman and Moldave (eds.) (1980) Meth. Enzymol. 65; Miller (ed) (1972) Exp. Mol. Genetics, Cold Spring Harbor, NY; Old and Primrose (1981) Principles of Gene Manipulation, Univ. of Cal. Press, Berkeley; Schlieff and Wensink (1982) Practical Methods in Molecular Biology; Glover (ed) 1985 (DNA Cloning, Vols. I and II, IRL Press, Oxford, UK; Sellow and Hollaender (1979) Genetic Engineering: Principles and Methods, Vols I, Plenum Press, NY; which are incorporated by reference in their entirety herein. Abbreviations, where employed, are those deemed standard in the field and commonly used in professional journals such as those cited herein.

The term "derived from", as it is used herein, refers to an actual or theoretical source or origin for isolated polynucleotides of the invention. For example, a polynucleotide that is "derived from" a particular polynucleotide (e.g., a GDF-9 gene) will be identical or highly homologous in nucleotide sequence to a relevant portion of the reference polynucleotide (e.g., a GDF-9 gene) molecule. Thus, for example, a polynucleotide that is "derived from" the first 10 kilobases or 3.3 kilobases of DNA immediately 5' of the coding sequence of a GDF-9 gene may correspond in nucleotide sequence to all or a portion of the 10 kilobase or 3.3 kilobase upstream sequence of wild-type GDF-9 gene. Isolated polynucleotides of the invention which are "derived from" GDF-9 genes (e.g., the region immediately 5' of the transcription initiation site) also include those which have been modified by insertion, deletion or substitution of one or more nucleotides but which retain substantially the same activity or function.

A DNA "coding sequence", "coding region", or a "sequence encoding" a particular protein is a DNA sequence which is transcribed and translated into a polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory

elements. The boundaries of the coding sequence are determined by a start codon at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian, animal, avian etc.) sources, and even synthetic DNA sequences. A transcription termination sequence will usually be located 3' to the coding sequence.

A "GDF-9 gene", as used herein, refers to a GDF-9 gene (e.g., a cloned genomic gene or a cDNA), including its untranscribed upstream and downstream regions and transcribed, untranslated regions from any species which naturally expresses GDF-9. For example, the nucleotide sequence for the genomic human GDF-9 gene is available at Genbank accession # AC004500. A "nonhuman GDF-9 gene", as used herein, refers to a GDF-9 gene (e.g., a cloned genomic gene or a cDNA), including its untranslated regulatory regions, from any species excepting human (e.g., avian, sheep, porcine, ovine, bovine). For example, the nucleotide sequence for the mouse GDF-9 gene is provided in U.S. Patent number 5,821,056, the complete contents of which are incorporated by reference herein.

The terms "regulatory element", "control element" and "regulatory sequence" are used interchangeably herein, and refer to a nucleic acid which, when operably linked to a gene, effects ("modulates") or causes transcription and/or expression levels of the gene in a cell. Such genetic regulatory elements are well known in the art and include, for example, promoters, enhancers and other cis-acting sequences involved in the binding of transcription factors/repressors. For example, a "promoter" is a regulatory element to which an RNA polymerase can bind, permitting the transcription and subsequent translation of an operatively linked coding sequence. Regulatory elements include both positive and negative ("repressors") regulators of transcription.

The term "promoter", as used herein, refers to a nucleotide sequence generally located adjacent to the 5' end of a structural gene that is involved in the initiation of transcription. Promoters contain DNA sequence elements that ensure proper binding and activation of an RNA polymerase, influence where transcription will start, and affect the level of transcription. The term "repressor" refers to a regulatory element generally located adjacent to the 5' end of a structural gene that is involved in the repression of the initiation of transcription, frequently by binding RNA polymerase, or

by binding factors interfering with the processivity or binding of RNA polymerase to the gene transcription start site. Further, specific regulatory sequences within or adjacent to promoters that are functional in the regulation (induction and repression) of gene expression responsive to stimuli or specific chemical species may also be present. The sizes of these regulatory elements are variable. In many cases, regulatory activity can be within approximately 500 bases to 3000 bases of sequence in the 5' direction (or upstream) to the site of transcription initiation. However, sequences out to approximately 4000 base pairs 5' to the structural gene have been implicated in the regulation of gene expression of certain genes. These control elements, however, may be located even further upstream of the structural gene under their control.

The majority of regulatory elements control initiation of transcription and processivity of the RNA polymerase in one direction only, so in order to be under the control of a regulatory element, a structural gene must usually be located downstream (in the 3' direction) of the element and in the correct orientation with respect to the element. The distance between the element and the structural gene is believed to be an important factor in gene expression level. One or several genes may be under the control of a single element or, conversely, one or more regulatory elements may control a single structural gene.

Environmental factors such as temperature, light, and oxygen tension, and chemical species such as nutrients, metabolites, heavy metal ions and steroids have been found to regulate gene expression. Factors that induce expression as well as factors that repress expression of genes have been identified. The exact mechanism of regulation by such signals or stimuli is likely to be complex, involving multiple protein interactions. By analogy to previous mechanistic studies of regulation, however, regulatory control is expected to involve changing the ability of RNA polymerase to bind to DNA sequences in the promoter region. One possible mechanism is the binding of regulatory protein to a DNA sequence at or near the position of binding of RNA polymerase to enhance or prevent transcription. A second possible mechanism is direct or indirect interaction of a signal (inducer or repressor) molecule with RNA polymerase, itself, to change its specificity for recognition and binding to a DNA sequence of the promoter. In either case, specific sequence(s) within the promoter would be involved in the mechanism of

regulation, the presence within a promoter region of one or several sequences can be important to the regulation of promoter activity.

"Expression" of a gene requires both transcription of DNA into mRNA and the subsequent translation of the mRNA into protein products.

5 The term "reporter gene", as used herein, refers to a gene encoding a protein which is readily quantifiable or observable. Because gene regulation usually occurs at the level of transcription, transcriptional regulation and promoter activity are often assayed by quantitation of gene products. For example, promoter regulation and activity has often been quantitatively studied by the fusion of the easily assayable *E. coli lacZ*
10 gene to heterologous promoters (Casadaban and Cohen (1980) J. Mol. Biol. 138:179-207). The structural gene for chloramphenicol acetyl transferase (CAT) and green fluorescence protein (GFP) are other genes commonly used to detect activity of a promoter or other regulatory sequence.

 The term "tissue-specific expression", as it is used herein, refers to a limited or
15 characteristic pattern of gene expression among cell types. In other words, expression of a gene is observed in certain tissues of an organism but not in other tissues. For example, "oocyte-specific" expression of a gene denotes that that gene is expressed in the oocyte and optionally limited other tissues, but is not expressed in all tissues (e.g., globally).

20 A "vector" is a replicon, such as a plasmid, phage, or cosmid, to which another DNA segment may be attached so as to bring about the replication of the attached segment.

 An "expression vector" means any DNA vector (e.g., a plasmid vector) containing the necessary genetic elements for expression of a desired gene, including a
25 promoter region of the present invention. These elements are "operably linked" to the gene, meaning that they are located at a position within the vector which enables them to have a functional effect on transcription of the gene. The regulatory elements need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be
30 present between a promoter and the coding sequence and the promoter can still be considered "operably linked" or "in operable linkage to" the coding sequence.

A cell has been "transformed" by exogenous DNA (e.g., a transgene) when such exogenous DNA has been introduced inside the cell membrane. Exogenous DNA may or may not be integrated into the chromosomal DNA comprising the genome of the cell. With respect to eukaryotic cells, though, a stably transformed cell is one in which the
5 exogenous DNA has become integrated into the chromosome such that it is inherited by daughter cells through chromosome replication.

A "host cell" is a cell that has been transformed, or is capable of transformation, by an exogenous nucleic acid molecule.

A "transgene" refers to a nucleic acid which is introduced into a cell. Typically,
10 the transgene is integrated into the genome of the cell following introduction. The transgene can encode a protein which is not expressed in the cell or which is expressed in the cell at low levels or in defective form.

A "transgenic animal" is an animal carrying in its cells at least one transgene. For example, the transgenic animal can contain in its cells a transgene corresponding to
15 a gene of another species which has been introduced into the germline of the animal, such that the introduced gene is present in all somatic and germline cells.

Identification and Isolation of GDF-9 Regulatory Sequences

GDF-9 regulatory sequences can be identified and isolated from a variety of
20 sources including mammalian and avian genomic and cDNA libraries. They can also be prepared (e.g., synthesized) from a variety of known and sequenced GDF-9 genes. For example, the nucleotide sequence for the genomic human GDF-9 gene is available at Genbank accession # AC007400. The nucleotide sequence for the ovine GDF-9 gene is published in Bodensteiner et al. (1999) *Biol. Reprod.* 60:381-386 and is accessible at
25 Genbank accession # AF078545. The nucleotide sequence for the mouse GDF-9 gene is published in Incerti et al. (1994) *Biochimica et Biophysica Acta* 122:125-128.

In one embodiment of the invention, GDF-9 regulatory sequences are obtained from the first 10 kilobases of DNA immediately 5' of the coding sequence of a GDF-9 gene, such as the mouse GDF-9 gene. In another embodiment, the GDF-9 regulatory
30 sequences are derived from the first 3.3 kilobases of DNA immediately 5' of the transcription initiation site of a GDF-9 gene. For example, the sequences can be derived from the 3.3 kilobase regulatory region from the mouse GDF-9 gene (mGDF-9) having

the nucleotide sequence of SEQ ID NO:1. In another embodiment, the GDF-9 regulatory sequences are derived from the first 300 base pairs of DNA immediately 5' of the transcription initiation site of a GDF-9 gene. In still another embodiment, the GDF-9 regulatory sequences are derived from the region from 3.3 kilobases to 10 kilobases 5' of the transcription initiation site of a GDF-9 gene. In another embodiment, the GDF-9 regulatory sequences are derived from the untranscribed 3' flanking region of a GDF-9 gene, or from a transcribed, untranslated region of a GDF-9 gene (e.g., an intron which is spliced out from the GDF-9 mRNA during processing).

Isolation and cloning of GDF-9 regulatory sequences can be performed using standard techniques well-known in the art, such as those discussed above and in the examples provided below.

The observed tissue-specific expression of growth differentiation factor 9 (GDF-9) demonstrates that there is one or more control element(s) for the GDF-9 gene which is permissive to GDF-9 expression in certain tissues, such as the oocyte (among other tissues) while it is repressive to this expression in other cell types.

As part of the present invention, it was discovered that the region immediately 5' of the transcription initiation site of the GDF-9 gene (spanning the first 10 kilobases) contains not only one or more regulatory elements (e.g., promoters) specifically directing the expression of GDF-9 in certain tissues, such as oocytes, testicular cells, and the hypothalamus, but also one or more repressor elements specifically inhibiting expression of GDF-9 in testicular cells. Accordingly, in one embodiment, the present invention provides isolated GDF-9 regulatory sequences derived from each of the first 10, 3.3, or 0.3 kilobases of DNA immediately 5' of the transcription initiation site of a nonhuman GDF-9 gene. Other embodiments include isolated polynucleotide molecules spanning each of the regions 3.3 kilobases, 10 kilobases, and 3.3-10 kilobases immediately 5' of the transcription initiation site of a nonhuman GDF-9 gene.

GDF-9 regulatory elements derived from these regions, and other untranscribed upstream and downstream regions, and transcribed, untranslated regions (e.g., introns) of GDF-9 genes, which regulate tissue-specific gene expression, can be identified by examining these regions for the presence of particular sequences which effect transcription and/or expression of a gene when the gene is operatively linked to the sequences. For example, it was found as part of the present invention that the

aforementioned 0.3, 3.3, and 10 kilobase regions from the region immediately 5' of the mouse GDF-9 transcription initiation site promote oocyte-specific expression of a reporter gene. It was also found that the 3.3 kilobase region promotes expression of a reporter gene in testis, but that the 10 kilobase region does not. This finding

5 demonstrates the presence of a testis-specific repressor element in the region from 3.3 to 10 kilobases immediately 5' of the transcription initiation site of the mouse GDF-9 gene. Further mapping of such promoter and repressor elements can be achieved by similarly testing smaller fragments of these regions to define the particular sequences involved in gene regulation.

10 Other functional mapping techniques also may be employed to further characterize and identify specific GDF-9 regulatory sequences. For example, nucleotide bases within these regions (e.g., the 0.3, 3.3 or 10 kilobase fragments of the 5' region) can be mutated by, for example site-directed mutagenesis, to add, delete or change one or more bases (e.g., 6-12 bases), followed by testing the mutated regulatory sequences
15 for activity either *in vitro* (e.g., by microinjection or transfection) or *in vivo* (e.g., in a transgenic animal) to see what functional effect the mutation had. From this information, nucleotide bases required for function (e.g., upregulation or downregulation) of the regulatory element can be determined. For example, if upon mutating a small (e.g., 6-12 base pair) segment of the 5' region of a GDF-9 gene,
20 transcription levels of a reporter gene operatively linked to the 5' region are diminished, then this small region can be concluded to be involved in promoting transcription (e.g., by binding to one or more transcription factors). Transcription factors that bind cis-acting regulatory elements may also interact with each other. Therefore, multiple reporter constructs can be developed and tested for the interaction between their
25 potential cis-elements and such binding proteins using the assays described herein.

Accordingly, in another embodiment, the invention provides a method for identifying GDF-9 regulatory elements by operatively linking various portions (e.g., overlapping portions) of the region immediately 5' of the transcription initiation site (e.g., the first 10 kilobases) of a GDF-9 gene, or other untranslated (e.g., 3') regions of a
30 GDF-9 gene, to a coding sequence (e.g., of a reporter gene), and comparing expression levels and patterns *in vivo* (e.g., in a transgenic animal) or *in vitro* (e.g., in an oocyte) among these constructs. Methods for preparing such constructs and for operatively

linking genes (e.g., within plasmid vectors) to such regulatory sequences are well known in the art. For example, a first expression construct which includes a portion of the 10 kilobase region immediately 5' of the transcription initiation site of a GDF-9 gene, in operative linkage with a reporter gene, can be introduced, e.g., by microinjection, into an oocyte or other tissue which naturally expresses GDF-9, followed by measurement of the activity of the reporter gene (e.g., by Northern blot). A second expression construct, including a different portion of the 10 kilobase region immediately 5' of the transcription initiation site of the same GDF-9 gene, also in operative linkage with the same reporter gene, can then be introduced into an oocyte, and activity of the reporter gene again measured. Upon comparison of the activities of the reporter genes controlled by the two different GDF-9 regulatory regions, a conclusion about the regulatory function (e.g., repressor or promoter activity) of the utilized portion of the region immediately 5' of the transcription initiation site can be drawn.

Alternatively, regulatory regions of GDF-9 genes can be identified by comparing untranscribed upstream and downstream regions, and transcribed, untranslated regions of a GDF-9 gene with 5' and 3' sequences from other known genes, for example, which are expressed in a similar tissue-specific pattern and, therefore, which may contain related or homologous regulatory elements. For example, as part of the present invention, the first 300 bases of the mouse GDF-9 gene immediately 5' of the transcription initiation site mouse GDF-9 gene was compared to the same region from the mouse ZP3 gene (Millar et al. (1991) *Molec. Cell. Biol.* 11:6197-6204), which is expressed predominantly in the oocyte, and homology was observed in certain regions. One such region is the E-box (CANNTG, where N is any nucleotide base) located approximately 200 bases upstream of the transcription start site. The E-box in the ZP3 gene is a conserved DNA element which has been shown to bind basic helix-loop-helix transcription factors, many of which are tissue-specific (Liang et al. (1997) *Development* 124:4939-4947).

Similarly, sequences derived from the untranscribed upstream and downstream regions, and transcribed, untranslated regions of a GDF-9 gene can be compared to sequences in a global gene database (e.g., GenBank) using computer programs well-known in the art to identify homologous regulatory sequences from other known genes.

In addition, DNA probes can be synthesized corresponding to a portion of an untranscribed or transcribed regulatory region of a GDF-9 gene, such as the mouse GDF-9 gene described herein, and used to screen for homologous sequences from other GDF-9 genes using, e.g., *in situ* hybridization. For example, DNA probes
5 corresponding to portions of first 3.3 kilobases of DNA (SEQ ID NO:1) immediately 5' of the mouse GDF-9 transcription start site can be generated and used to screen for homologous sequences which hybridize to the probes, preferably under high stringency conditions.

Accordingly, GDF-9 regulatory sequences of the invention include homologous
10 sequences from other species and sequences which have been altered but retain a high percentage of sequence identity. Preferred GDF-9 regulatory sequences have at least about 50% homology, more preferably at least about 60% homology and most preferably at least about 70-99 % homology with regulatory sequences derived from the first 10, 3.3 or 0.3 kilobases of DNA upstream of the mouse GDF-9 gene transcription
15 start site. Homology refers to sequence similarity between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base, then the molecules are homologous at that position. A degree of homology between sequences is a function of the number of matching or
20 homologous positions shared by the sequences.

Using the aforementioned methods, tissue-specific regulatory elements derived from GDF-9 genes can be identified and then synthesized, cloned or otherwise isolated. Accordingly, other embodiments of the invention include isolated tissue-specific (e.g., oocyte-specific or testis-specific) promoter and enhancer elements derived from a GDF-
25 9 gene. Other embodiments include isolated tissue-specific repressor elements derived from a GDF-9 gene. These regulatory elements can be operatively linked to a gene of interest, directing or repressing expression of the gene in a tissue-specific manner. For example, repression of certain functionally active genes in the testis is desirable, and the GDF-9 repressor element has the ability to direct the repression of such genes in
30 testicular cells.

Expression Vectors and Transfection of Cells with GDF-9 Regulatory Sequences

In another embodiment, the invention provides an expression vector which includes one or more of the aforementioned GDF-9 regulatory elements in operative linkage with a gene of interest, such that expression of this gene is under control of the GDF-9 regulatory element. In one embodiment, the regulatory element is derived from the 10 kilobases of the region immediately 5' of the transcription initiation site of a GDF-9 gene. Any gene that expresses a protein of interest may be employed in the methods of the invention. Similarly, the invention may be used to repress expression of any gene desired to be downregulated in a specific tissue. Examples of such genes include GDF-9 itself, growth factors, genes encoding antigens to tumors or pathogens, and reporter genes. However, this is not intended to be a limiting list.

When employed in the context of heterologous structural genes, the precise optimal location of the control sequences of the invention with respect to the transcription initiation site may vary. Generally, the control sequences are placed up to about 300 nucleotides or more from a transcription initiation site. However, in other embodiments, control sequences are located within 150 nucleotides of the transcription initiation site. Therefore, to employ regulatory elements of the invention in the context of a heterologous gene, the regulatory elements can simply be inserted into an expression construct upstream of a transcription initiation site. Additionally, as is known in the art, it is generally desirable to include TATA-box sequences upstream of and proximal to a transcription initiation site of the heterologous structural gene. Such sequences may be synthesized and inserted in the same manner as the GDF-9 control sequences. Alternatively, one may desire to simply employ the TATA sequences normally associated with the heterologous gene. Such TATA sequences are most desirably located between about 20 and 30 nucleotides upstream of transcription initiation.

GDF-9 regulatory sequences of the present invention, whether positive, negative, or both, may be cloned into such expression vectors in the form of multiple units, in numerous various combinations and organizations, in forward or reverse orientations, and the like. Moreover, in the context of multiple unit embodiments and/or in embodiments which incorporate both positive and negative control units, there is no requirement that such units be arranged in an adjacent head-to-head or head-to-tail

construction in that the improved regulation capability of such multiple units is conferred virtually independent of the location of such multiple sequences with respect to each other. There is no requirement that each unit comprise the same positive or negative element. All that is required is that such sequences be located upstream of and sufficiently proximal to a transcription initiation site for the gene of interest.

The control sequences of the present invention may be beneficially employed in the context of any heterologous gene, with or without additional homologous or heterologous control or promotion sequences. The present invention encompasses GDF-9 gene promoters and other regulatory sequences which function in the induction or repression of GDF-9 expression in response to tissue-specific factors (e.g., transcription factors) that induce and repress GDF-9 expression. A GDF-9 promoter-GFP reporter construct was used in the instant invention, however any suitable reporter gene may be used to measure the activity of the GDF-9 promoter.

In yet another embodiment, the invention provides a cell, preferably a cell which naturally expresses GDF-9 (e.g., an oocyte cell line), stably transformed with one or more of the aforementioned GDF-9 5'-flanking region polynucleotides, or with one of the expression vectors previously described. For example, the cell can be transformed *in vitro* or *in vivo* with the isolated GDF-9 regulatory region which, in certain embodiments, can be inserted into a particular locus of the cells genome. Methods for such site-specific insertion of transgenes into chromosomal DNA are well known in the art and include, for example, the mammalian Cre/lox system (Sauer et al. (1998) *Methods* 14:381-392) or homologous recombination (see e.g., 5,614,396). By targeting the regulatory sequences to locations upstream of endogenous genes, expression of these genes can be controlled accordingly (e.g., upregulated or downregulated).

Alternatively, the GDF-9 control elements of the invention can be introduced into cells in operational linkage with a transgene. Introduction of the GDF-9-transgene construct into a host cell will result in expression or repression of the transgene, depending on the element or combination of elements contained within the construct. In specific embodiments, this expression or repression occurs specifically in oocytes or testicular cells. For example, an expression vector containing the aforementioned 3.3 kilobase mouse GDF-9 promoter element in operative linkage with a gene encoding luciferase will cause oocyte and testis-specific expression of luciferase upon

introduction into oocytes or testicular cells. Similarly, transformation of a testicular cell with an expression vector containing the previously described 3.3 to 10 kilobase region from the mouse GDF-9 5' UTR, or derivatives thereof, in operative linkage with the luciferase gene will cause testis-specific repression of the luciferase gene.

5 Another embodiment of the present invention provides the utilization of the previously described GDF-9 regulatory elements to identify proteins and molecules involved in the regulation of GDF-9 transcription and translation. In one embodiment, one or more GDF-9 regulatory elements, such as a GDF-9 promoter, is operatively linked to a reporter gene, such as the green fluorescent protein gene. Factors or
10 compounds can then be tested using this construct for their ability to effect expression levels of the reporter gene compared to a control. Such screening methods and how to perform them are well known to those of ordinary skill in the art. Suitable reporter assays include, for example, CAT assays and luciferase assays. The invention is not restricted to these suggested screening methods. Any method known in the art may be
15 employed whereby compounds can be tested for their ability to increase or stimulate the expression of a reporter gene compared to that of a control reporter gene, (e.g., the SV40 β -galactosidase reporter gene). Effects of the test compound are determined by changes in reporter gene activity.

Transcription factors that bind to these specific regulatory sequences can be
20 characterized using gel mobility shift assays and these transcription factors can be cloned using these specific sequences as probes in screening expression libraries. After these transcription factors have been identified, they too, may be used as targets for identifying other inhibitors. For example, these transcription factors can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Patent No.
25 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Biotechniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with the transcription factors. Such transcription factor-binding proteins are also likely to act as modulators, e.g., inhibitors of GDF-9 expression.

Uses of the Invention

GDF-9 regulatory sequences of the present invention can be used to selectively express GDF-9 and other gene products of interest in the oocyte, testis, hypothalamus, placenta and other tissues in which GDF-9 mRNA is naturally transcribed. GDF-9 regulatory sequences may also be used to repress GDF-9 expression and other gene products (e.g., tumor antigens) in specific tissues, such as the testis.

Tissue-specific expression of genes, particularly in the oocyte, can be used to develop treatments for infertility, or conversely, can be used for purposes of contraception, either through the selective expression of a nonfertility factor in oocytes, or through the selective repression of a fertility factor in testicular cells.

GDF-9 regulatory sequences of the invention also can be used to develop transgenic animals in which one or more genes are specifically expressed or repressed in the oocyte, testis, and hypothalamus. Techniques for generating such transgenic animals are now well known in the art.

As previously discussed, GDF-9 regulatory sequences of the present invention also can be used to identify other homologous tissue-specific regulatory elements. Accordingly, an important aspect of this invention is the utility of the disclosed elements and genetic regions as probes for the detection of similar regulatory sequences and regions in other genomes, or elsewhere within a given genome.

GDF-9 regulatory sequences of the invention also can be used to selectively express or repress desired gene products in germ cells. For example, the Cre/lox site-specific recombination system previously described herein can be used in oocytes (e.g., from mouse), or in both male and female germ cells (e.g., from mouse), to target GDF-9 repressor elements upstream of selected genes so that they inhibit expression of the genes, thereby creating germ cell-specific knockouts. This ability is particularly important for genes that cause early death in conventional knockouts, since the reproductive function of the eliminated genes cannot be assessed in that situation.

As previously discussed, GDF-9 regulatory sequences of the invention also can be used to identify and to refine germ-cell promoter elements. Small polynucleotide molecules based upon DNA sequence from identified GDF-9 promoter region(s) can be labeled and used as probes for the identification of other regions similar in sequence in a

given genome. The functionality of such regions as promoters of expression in germ cells can be assessed using methods disclosed in this invention.

GDF-9 regulatory sequences of the invention also can be used as reagents for the further identification of testicular cell repressor elements. For example, small
5 polynucleotide molecules based upon DNA sequence from the 3.3 kb-10 kb GDF-9 testis-specific repressor region discussed herein can be labeled and used as probes for the identification of other regions similar in sequence in a given genome. The functionality of such regions as repressors in the testis can be assessed using methods disclosed in this invention. GDF-9 regulatory sequences of the invention also can be
10 used to identify somatic cell repressor elements that are similar to the aforementioned testicular cell repressor element for the GDF-9 gene.

GDF-9 regulatory sequences of the invention also can be used as reagents for use in gene therapy. For example, GDF-9 promoter and repressor elements can be included in somatic gene therapy vectors. In the case of the GDF-9 testis-specific repressor
15 element, inclusion in such a vector permits the expression of a desired protein (such as a toxin or a drug) in all tissues in which the vector was incorporated (generally any rapidly dividing cell) with the exception of testicular cells. This permits the male germline cells to survive treatments of this kind, thus protecting male fertility.

GDF-9 regulatory sequences of the invention also can be used to generate
20 "marked" germ cells. By permitting the directed expression of a reporter gene specifically to certain tissues, such as oocytes or testicular cells, the present invention permits the generation of germline cells including a recognizable or measurable protein marker. For example, expression of green fluorescent protein in oocytes under the direction of the GDF-9 promoter results in a population of oocytes which each contain
25 the characteristic fluorescent protein. Other such reporter genes are well known in the art. Such cell lines can be used for testicular or ovarian chimeric analysis, or for germ cell kinetic studies. Also, they could be easily separated from somatic cells by flow-sorting technology without the necessity of antibody staining, permitting the generation of pure germ cell populations for library construction, in vitro culture or fertilization.

This invention is further illustrated by the following examples that should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

5 **Exemplification**

1. Cloning of the GDF-9 genomic locus

A GDF-9 cDNA probe was amplified using PCR primers based on the human GDF-9 published sequence in a PCR reaction from mouse ovary RNA. This mouse cDNA probe was used to screen a mouse genomic λ FIX II library from mouse 129SvEv
10 (Stratagene). Seven independent phages were identified, purified and screened by Southern blot analysis. Two of the GDF-9 genes containing phage were subcloned and further characterized. One of these two phage, B/S4, contained 10kb of sequence 5' to the beginning of GDF-9 exon I, GDF-9's entire coding region GDF-0 exon I (387bp), intron (2.9kb), and exon II (909bp), and - 1kb of 3' genomic sequence (Incerti et al.
15 1994).

2. Construction of GDF-9 transgenes

a) Alteration of the GDF-9 initiation codon to a BamHI site

A 700bp PstI fragment which contains -400 base pairs of GDF-9 exon I was subcloned into pBluescript SK+ (Stratagene). Two polymerase chain reactions (PCR)
20 were performed on this subclone using one vector specific primer and one primer that matched the region of GDF-9's initiation codon with several mismatches. Primer sequences were as follows (mismatched bases are underlined new BamHI site is in italics): GDF9-E1: 5'TTCTTC AAG GGATCC CACTTTCCCAG3' (SEQ ID NO:4) and GDF9-E2: 5' CTGGGAAGTG GGATCC CTGGAAGAA3'(SEQ ID NO:5).
25

b) Introduction of mutated GDF-9 regulatory region into pE22

The new BamHI site was used to reassemble the 700 base pair region from the two PCR products. This region was sequenced to ensure that there were no PCR errors. In a several step process, the full 15.5 kilobase genomic locus initially cloned in B/S4
30 was reassembled to include the unique BamHI site and mutated initiation codon (plasmid pE22).

c) Construction of GDF-9 Transgene 1 (10kb-GDF-9-GFP) (see Figure 2)

Plasmid pE22 was linearized with BamHI and the 700 base pair coding region of green fluorescent protein was inserted. A -17.5 kilobase fragment (SalI/NotI) was injected into pronuclei of fertilized 1-cell embryos. Mice in which transgene integration
5 had occurred were identified by Southern blot analysis of tail DNA using both a GFP probe and a GDF-9 exon II probe. Two transgene-positive founders were generated.

d) Construction of GDF-9 Transgene 2 (3.3kb-GDF-9-GFP) (See Figure 2)

Plasmid pE22 was cut with EcoRV and BamHI to generate a 3.3 kilobase piece
10 immediately 5' to the initiation codon. This was ligated immediately upstream of the ATC of enhanced green fluorescent protein cDNA (Clontech) containing its own SV40 polyadenylation sequence. A -4.0 kilobase fragment was injected as above and 6 independent transgene-positive lines were identified using the eGFP coding region probe.

15

3. Transgene Expression Analysis

Total RNA was isolated from brain, heart, lung, stomach, liver, spleen, kidney, small intestine, skin, uterus, ovary and testes using RNA Stat 60 (Leedo Medical, Houston, TX) according to manufacturer's instructions, from each transgene-positive
20 line. RNA was electrophoresed on a 1.2% agarose/formaldehyde gel, transferred to nylon membrane, probed with GFP or eGFP coding region ³²P-labeled probe and analyzed autoradiographically as previously described (Dube et al, 1998).

Cellular localization of transgene expression was conducted by in situ hybridization. Briefly, ovaries from transgene-positive mice and transgene-negative (wild-type) mice
25 were fixed overnight in 4% paraformaldehyde (PFA), processed and embedded in paraffin. 5 micron thick sections were cut, dewaxed, rehydrated and post-fixed in 4% PFA. Sections were pretreated in 50µg/ml proteinase K and acetic anhydride in 0.1M triethanolamine. Sections were hybridized overnight at 55°C with either ³⁵S labeled sense or antisense riboprobe for FGP or eGFP. Sections were washed in 5XSSC/50%
30 formamide and 0.1XSSC at 65°C. Slides were dipped in NTBZ liquid photographic emulsion (Kodak) and exposed at 40°C for 3-10 days. After development, signal was visualized by dark field microscopy.

Of the two lines of transgenic mice incorporating the first GDF-9-GFP construct (containing 10 kb of the 5' flanking sequence), *in situ* hybridization and multi-tissue northern blot analysis of mouse tissues demonstrated marked ovary-specific expression in the first line (Figure 3). The presence of signal in other tissues in the second line was most likely due to elements present at the site of transgene integration.

Six lines of mice carrying the second GDF-9-GFP construct were established. This construct contains only 3.3 kb of 5' flanking sequence linked to an enhanced version of GFP. Two of the six mouse lines showed oocyte-specific expression in the ovary by *in situ* hybridization. By multi-tissue northern blot analysis, signal was detected in the ovary, and to a much higher level in the testis (Figures 4 and 5). *In situ* hybridization for eGFP in the testis demonstrated that expression was restricted to the germ cells, specifically, primary spermatocytes through round spermatid stages. Similar studies in the ovary determined that expression is limited to the oocytes. One of the two lines also had a strong signal in the lung, most likely attributable to integration of the transgene near a strong lung-specific promoter or enhancer.

4. Sequence analysis of GDF-9 upstream region

The 3.3kb region present in GDF-9 transgene 2 was subcloned (Figure 7) and sequenced with 1X-3X coverage with vector primers by a standard dye termination method using fluorescently labeled dideoxy-nucleotide triphosphates and an ABI automatic sequence. Raw sequences were assembled using Lasergene and DNA Strider sequence analysis software. The complete sequence is shown in SEQ ID NO:1. Potential transcription factor binding sites and regulatory sequences were identified using Signal Scan (available through the University of Singapore home page). The human genome sequencing project sequenced the human GDF-9 within a much larger P1 clone from human chromosome 5 (Genbank accession # AC004500). Regions of sequence identity between human and mouse 5' sequences (See Figure 1) were revealed using a local alignment program LFASTA and LALIGN (GCG sequence analysis software). Promoter regions of other oocyte-expressed genes were scanned manually for interesting transcription factor binding sites conserved between mouse and human GDF-9 (particularly GATA-1 and E-boxes). To identify potentially relevant sequences a comparison with the 5' flanking region of the mouse zona pellucida 3 (mZP3) gene was

made. Transgene studies using just the first 280 bp of 5' flanking sequence from mZP3 demonstrated that this region is sufficient to direct oocyte expression. Sites shown to be critical for mZP3 oocyte expression, including the E-box, are also found within the first 300bp of 5' flanking sequence from mGDF-9 (see Figure 6).

5

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following

10 claims.

MTN-029PC

10/018638

JG13 Rec'd PCT/PTO 19 DEC 2001

- 23 -

What is claimed is:

1. An isolated polynucleotide comprising a GDF-9 regulatory element derived from a region of a nonhuman GDF-9 gene selected from the group consisting of the first 10
5 kilobases of DNA immediately 5' of the transcription start site, an intron, and the first 1 kilobase of DNA immediately 3' of the transcription termination site, wherein said isolated polynucleotide is greater than 261 nucleotides in length.
2. The polynucleotide of claim 1 wherein the regulatory element is derived from
10 the first 3.3 kilobases of DNA immediately 5' of the transcription start site of the nonhuman GDF-9 gene.
3. The polynucleotide of claim 1 wherein the regulatory element is derived from
15 the first 300 base pairs of DNA immediately 5' of the transcription start site of the nonhuman GDF-9 gene.
4. An isolated polynucleotide comprising the first 10 kilobases of DNA immediately 5' of the transcription start site of a nonhuman GDF-9 gene.
- 20 5. An isolated polynucleotide comprising the first 3.3 kilobases of DNA immediately 5' of the transcription start site of a nonhuman GDF-9 gene.
6. An isolated polynucleotide comprising the region from 3.3 kilobases to 10 kilobases immediately 5' of the transcription start site of a nonhuman GDF-9 gene.
25
7. An isolated oocyte-specific regulatory element derived from the 10 kilobases of DNA immediately 5' of the transcription start site of a GDF-9 gene, wherein said oocyte-specific regulatory element is greater than 261 nucleotides in length.
- 30 8. An isolated testis-specific regulatory element derived from the 10 kilobases of DNA immediately 5' of the transcription start site of a GDF-9 gene, wherein said testis-specific regulatory element is greater than 261 nucleotides in length.

SUBSTITUTE PAGE

AMENDED SHEET

MTN-029PC

- 24 -

9. The regulatory element of claim 8, wherein said element is derived from the first 3.3 kilobases of DNA immediately 5' of the transcription start site of a GDF-9 gene, and wherein said element causes tissue-specific expression of a gene operatively linked to the element in the testis.
10. The regulatory element of claim 8, wherein said element is derived from the region from 3.3 kilobases to 10 kilobases of DNA immediately 5' of the transcription start site of a GDF-9 gene, and wherein said element downregulates expression of a gene operatively linked to the element in the testis.
11. An expression vector comprising the isolated GDF-9 polynucleotide of any one of claims 1, 4, 5 or 6 operably linked to a gene.
12. The expression vector of claim 11, wherein the gene is a reporter gene.
13. An oocyte containing the polynucleotide of any one of claims 1, 4, 5 or 6.
14. A method for obtaining oocyte-specific expression of a gene, the method comprising transfecting an oocyte with the isolated polynucleotide of claim 1.
15. The method of claim 14, wherein said polynucleotide is operably linked to a gene.
16. A method for obtaining testis-specific expression of a gene, the method comprising transfecting a testicular cell with the isolated polynucleotide of claim 2.
17. The method of claim 16, wherein said polynucleotide is operably linked to a gene.
18. A method for down-regulating the expression of a gene in the testis, comprising transfecting a testicular cell with the isolated polynucleotide of claim 4.

SUBSTITUTE PAGE

AMENDED SHEET

MTN-029PC

- 25 -

19. A method for identifying tissue-specific regulatory elements for GDF-9 expression comprising, in any order, the steps of:

5 a) introducing into a cell a first expression vector comprising a portion of the region spanning 1 to 10 kilobases immediately 5' of the transcription start site of a GDF-9 gene;

b) introducing into a cell a second expression vector comprising a portion of the region spanning 1 to 10 kilobases immediately 5' of the transcription start site of a GDF-9 gene, wherein the portion differs from that contained in said first expression vector;
10 and

c) comparing expression patterns of said first and second vectors.

20. The method of claim 19, wherein said cell from an oocyte.

15 21. The method of claim 19, wherein said expression constructs are introduced into said cell via microinjection.

22. The method of claim 19, wherein said expression constructs are introduced into said cell via injection of a transgenic animal.

20

SUBSTITUTE PAGE

AMENDED SHEET

CORRECTED
VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/85, A61K 48/00, C12N 15/11	A3	(11) International Publication Number: WO 99/50406 (43) International Publication Date: 7 October 1999 (07.10.99)
(21) International Application Number: PCT/US99/07185 (22) International Filing Date: 31 March 1999 (31.03.99) (30) Priority Data: 60/080,108 1 April 1998 (01.04.98) US (71) Applicant (for all designated States except US): METAMORPHIX, INC. [US/US]; 1450 South Rolling Road, Baltimore, MD 21227 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): MATZUK, Martin, Matthew [US/US]; 2926 Russett Place West, Pearland, TX 77584 (US). ELVIN, Julia, Andrea [US/US]; 3114 Stanton Street, Houston, TX 77025 (US). (74) Agents: MANDRAGOURAS, Amy, E. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i> (88) Date of publication of the international search report: 18 November 1999 (18.11.99)
(54) Title: GROWTH DIFFERENTIATION FACTOR-9 REGULATORY SEQUENCES AND USES THEREFOR (57) Abstract Isolated GDF-9 regulatory sequences are disclosed, as well as methods of using the sequences to modulate tissue-specific expression of genes. The GDF-9 regulatory sequences include, for example, enhancer and promoter elements that naturally drive transcription of GDF-9 in specific tissues. The GDF-9 regulatory sequences can be derived from the untranscribed upstream (e.g., first 10 kilobases) and downstream regions, and transcribed, untranslated regions of a GDF-9 gene.		

*(Referred to in PCT Gazette No. 17/2000, Section II)

1/7

71.3% identity in 376 nt overlap; init: 76, opt: 608

```
      10      20      30      40      50      60
-mGDF GTTGAAGAACTAAGAAGGAACGCATCTCCTCGTGCCAGTCTTCTTTCTAAAGGTCGGGG
      :::::  :::::  :::::  :::::  :::::  :::::
-hGDF CTTGGGAGAACTAGTGAGGAACATATTTCTCCATGCCAGTCTTCTTCTAAAGGCCAGGA
      10      20      30      40      50      60

      70      80      90     100     110
-mGDF AGGCCCTGCGGTGGTC-CTTAAATAAAATCTAGGTGT---GGAAGACTGACTTCACTTCT
      ::  ::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
-hGDF AGAGCCTAGCTTGGTCTCTTAAATAAAATTTTCAGGTGTGTGGGTGGACTACGGTCACTTCT
      70      80      90     100     110     120

                                GATA-1      E-Box
      120     130     140     150     160     170
-mGDF GTAATCAGACCTCAAGCATGCCCTAGCTGTACATTGTTTCTTATCAGCTCTACAGCAGTCA
      :::::  :::::  :::::  :::::  :::::  X:::::  :::::  :::::
-hGDF GTAATCAGGCCTCAAGTATGCCCTAGCTGAAGATT---TTATCAGCTCTATATCAAGCA
      130     140     150     160     170

                                Oct-1
      180     190     200     210     220     230
-mGDF GCTGACAACACGGT+TTTATCCAATTTGTTCTCATTAGATACAGATTATCTACTTATTT
      ::::X :::::  :::::  :::::  :::::  :::::  :::::  :::::
-hGDF GCTGATAACACCTTATTTAGCCAATTTGT--TAATTAGATACAGATTACTTGGTTATTT
      180     190     200     210     220     230

      240     250     260     270     280     290
-mGDF GGTCTTCTTTCCCTGTTTCTCCTGGACATTCTATTTCAAACCCACCAAGTATCGTCACCCCT
      :::::  ::  ::  :::::  :::::  :::::  :::::  ::  ::  ::
-hGDF AGCTTTC-CTCTCT-TTTCCTCTGG-CATTTACGTAAAACCC----GTTACTGTTACACT
      240     250     260     270     280

      300     310     320     330     340     350
-mGDF AGTGGACTACTTATATAGGTGAGAGCCCAACAGAAACCCAAGGA-ATTATAAAACAAAAC
      :  :::::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
-hGDF ACCGGACTAACTATGTAGCTGAAAGTCTAAGATAAATTGAGGAATGATCTTAAATAAAAT
      290     300     310     320     330     340

      360     370
-mGDF CTCTTAGGGCTGGAGA
      ::::  :  ::::
-hGDF CTCTCATTTTGTGAGA
      350     360
```

5'

Figure 1

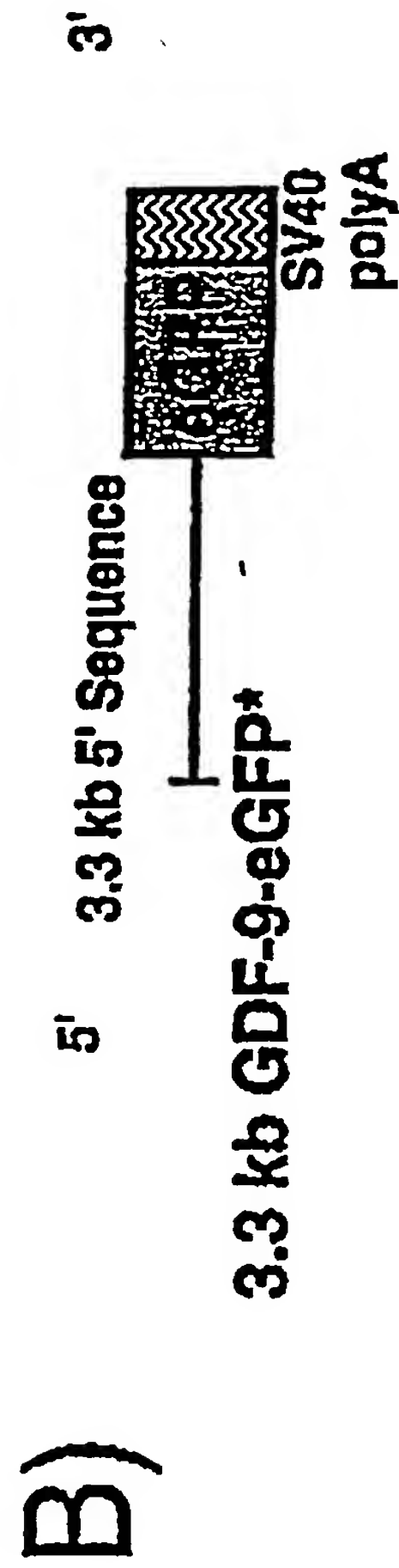
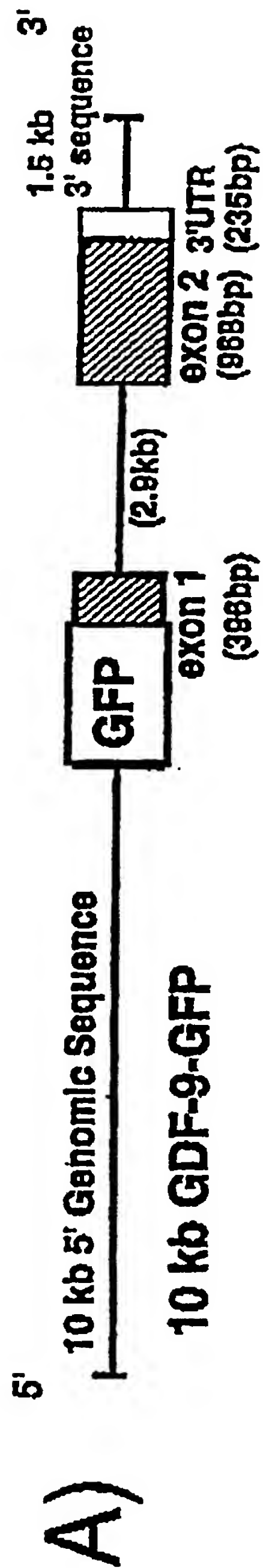


Figure 2

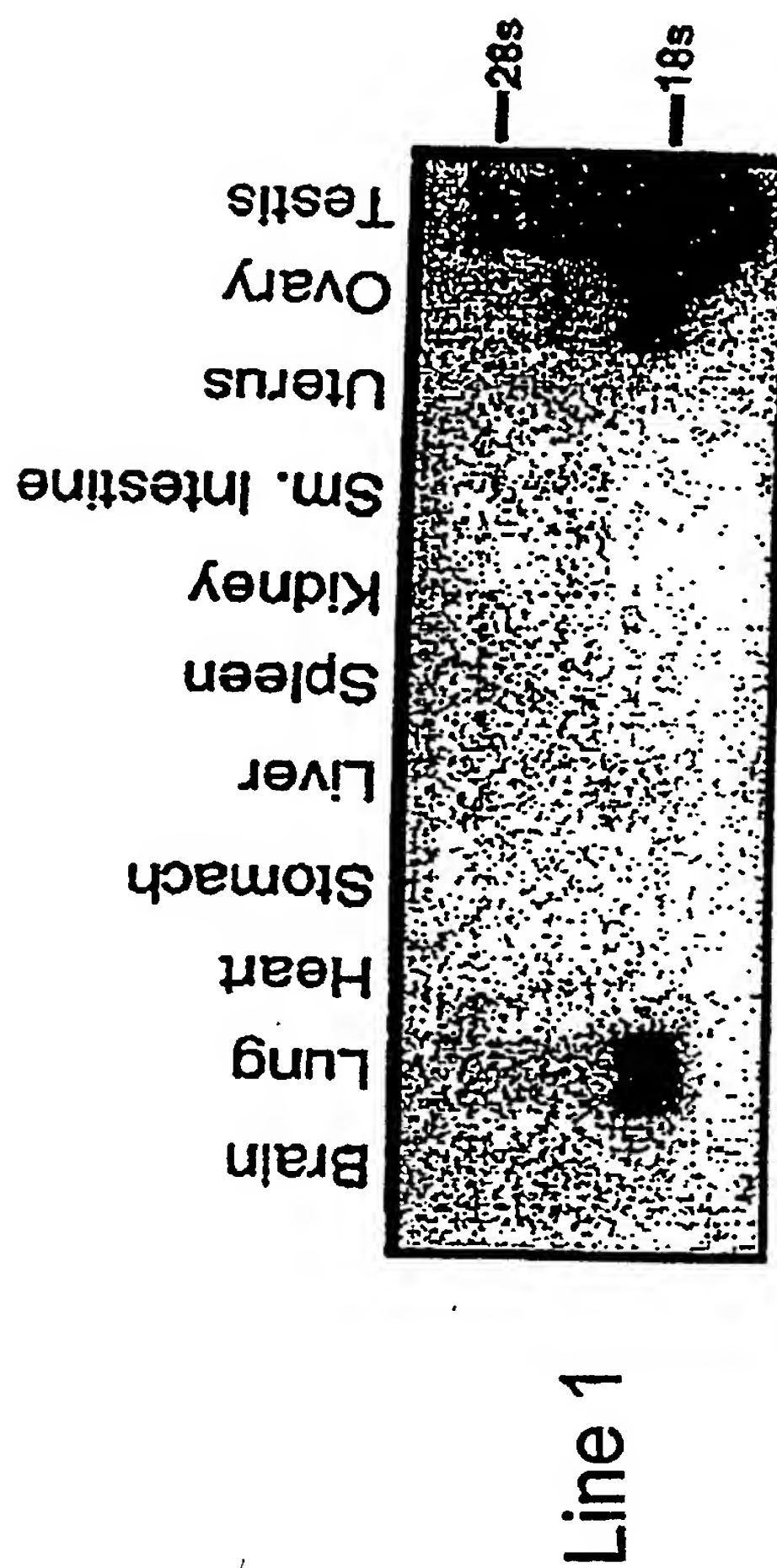


Figure 3

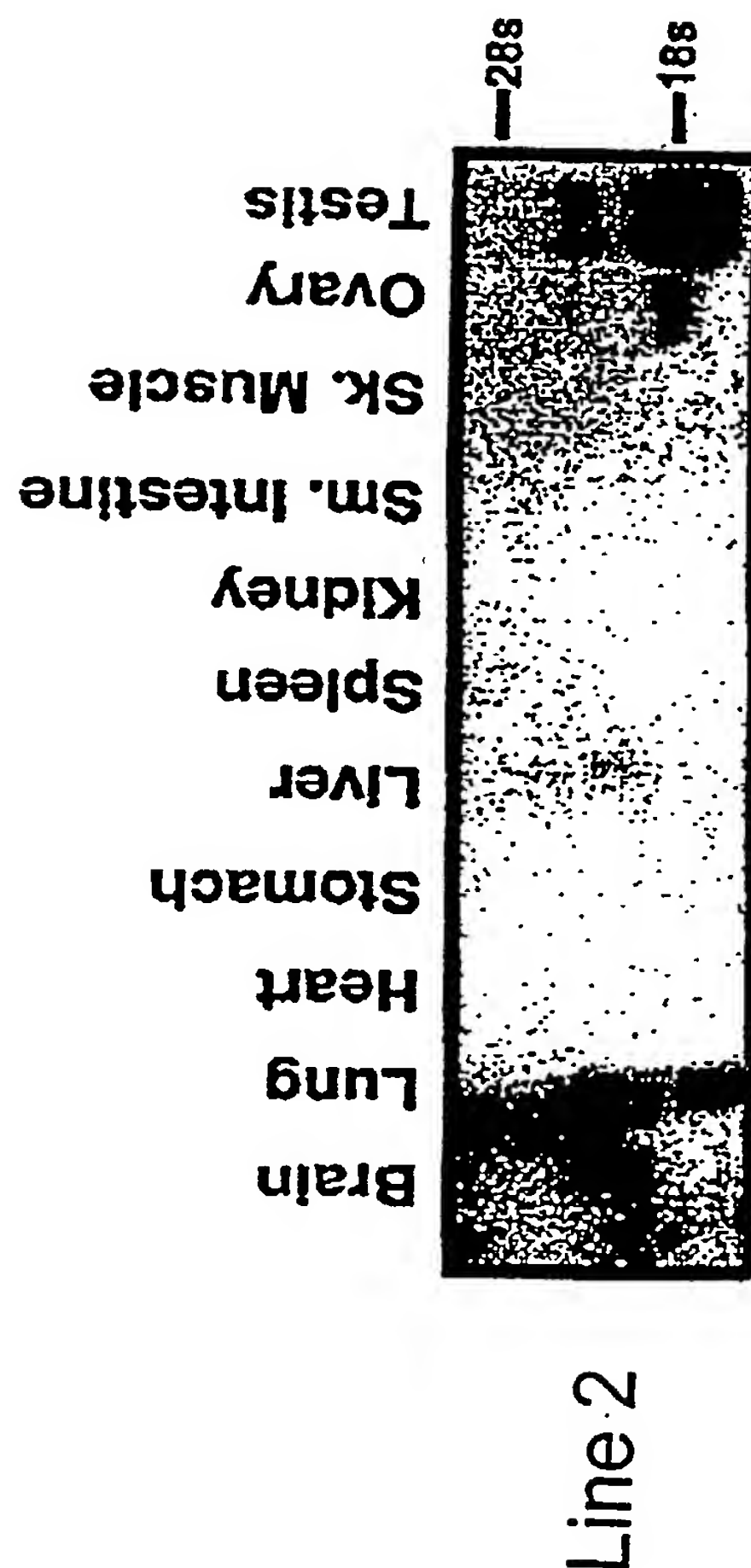
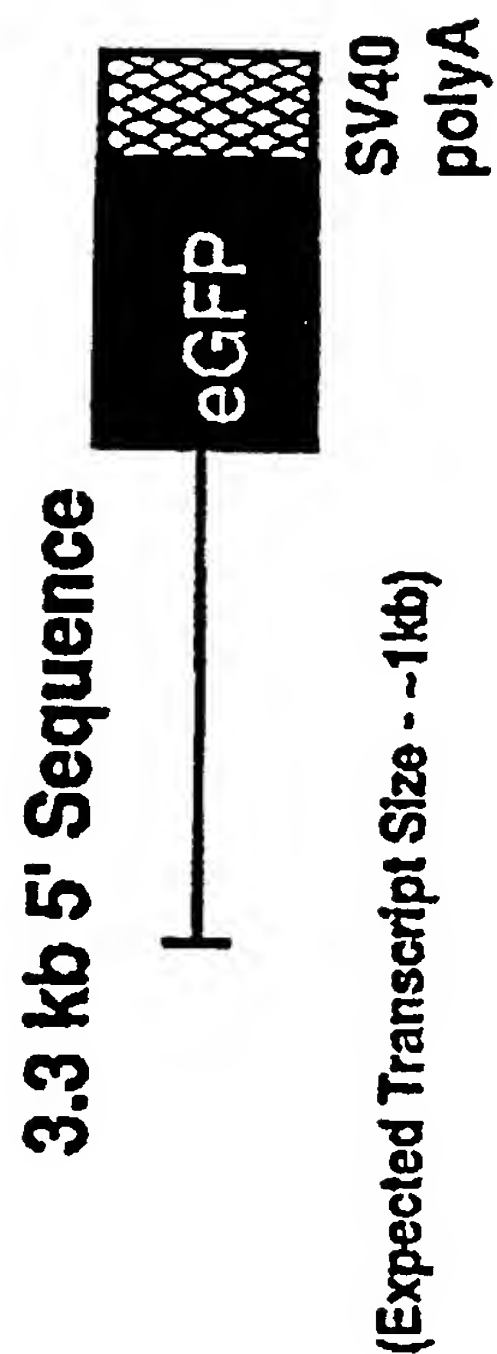


Figure 4

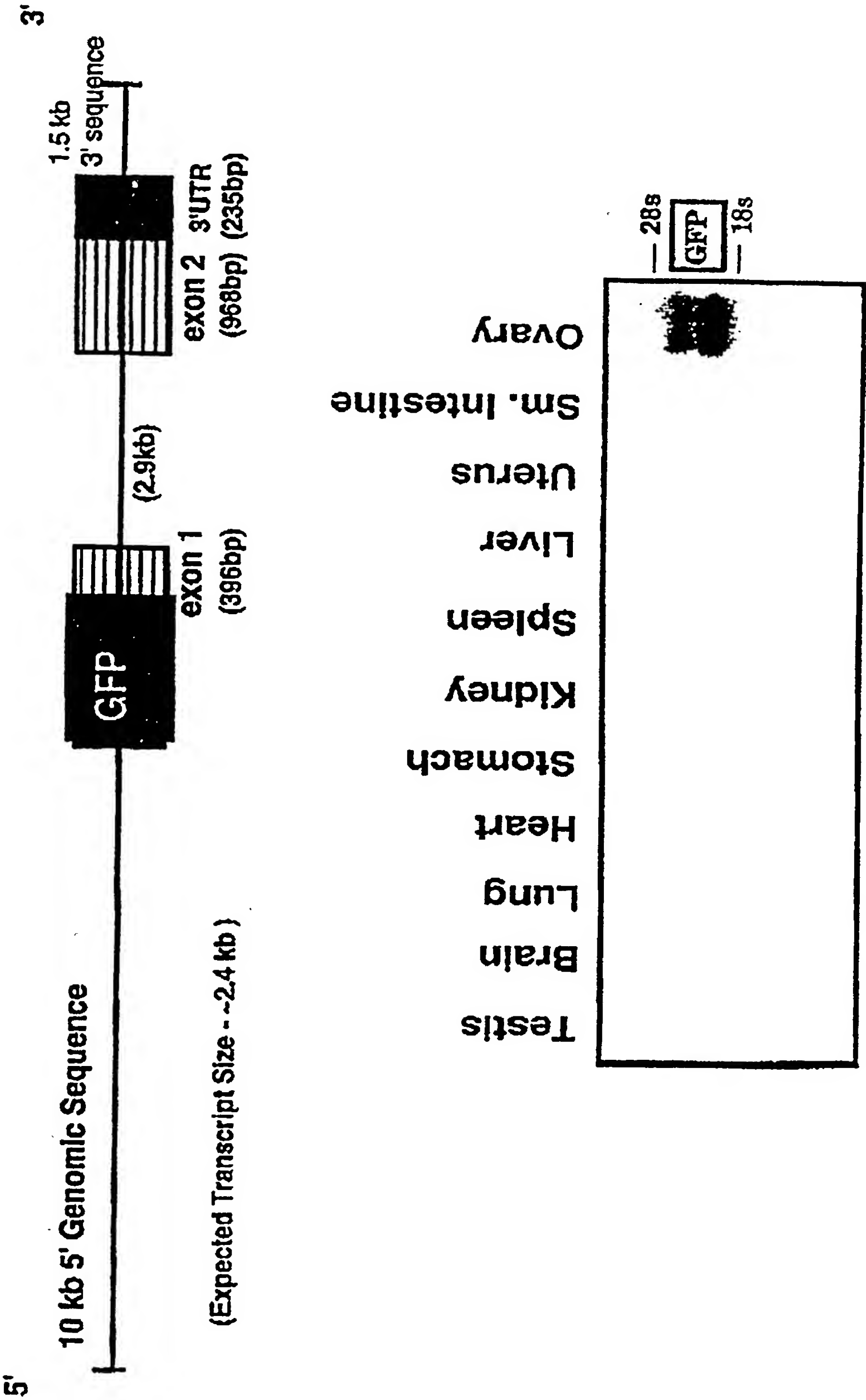
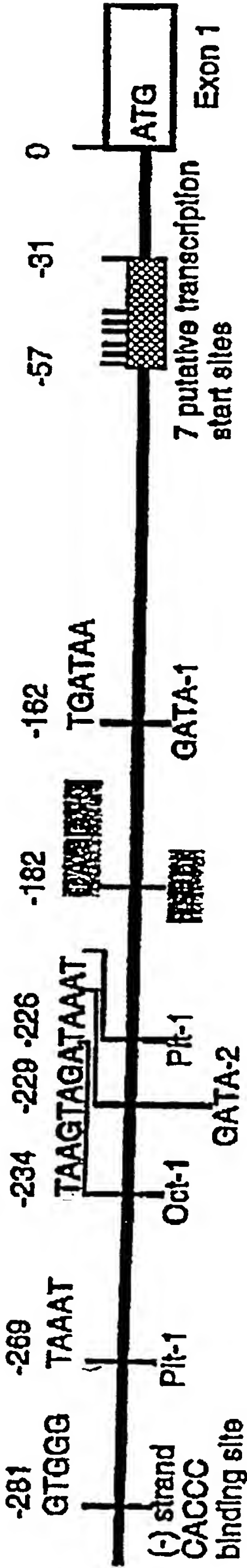


Figure 5

mGDF-9 5' Flanking Region



mZP3 5' Flanking Region *

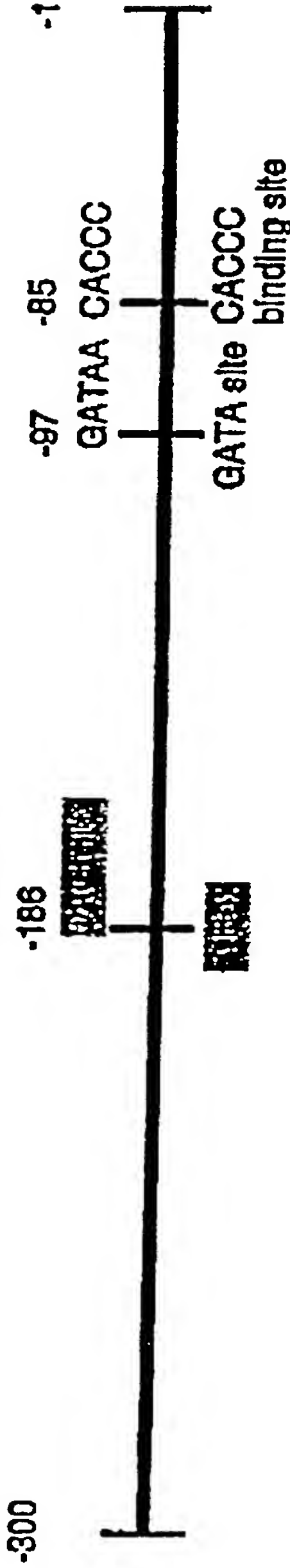


Figure 6

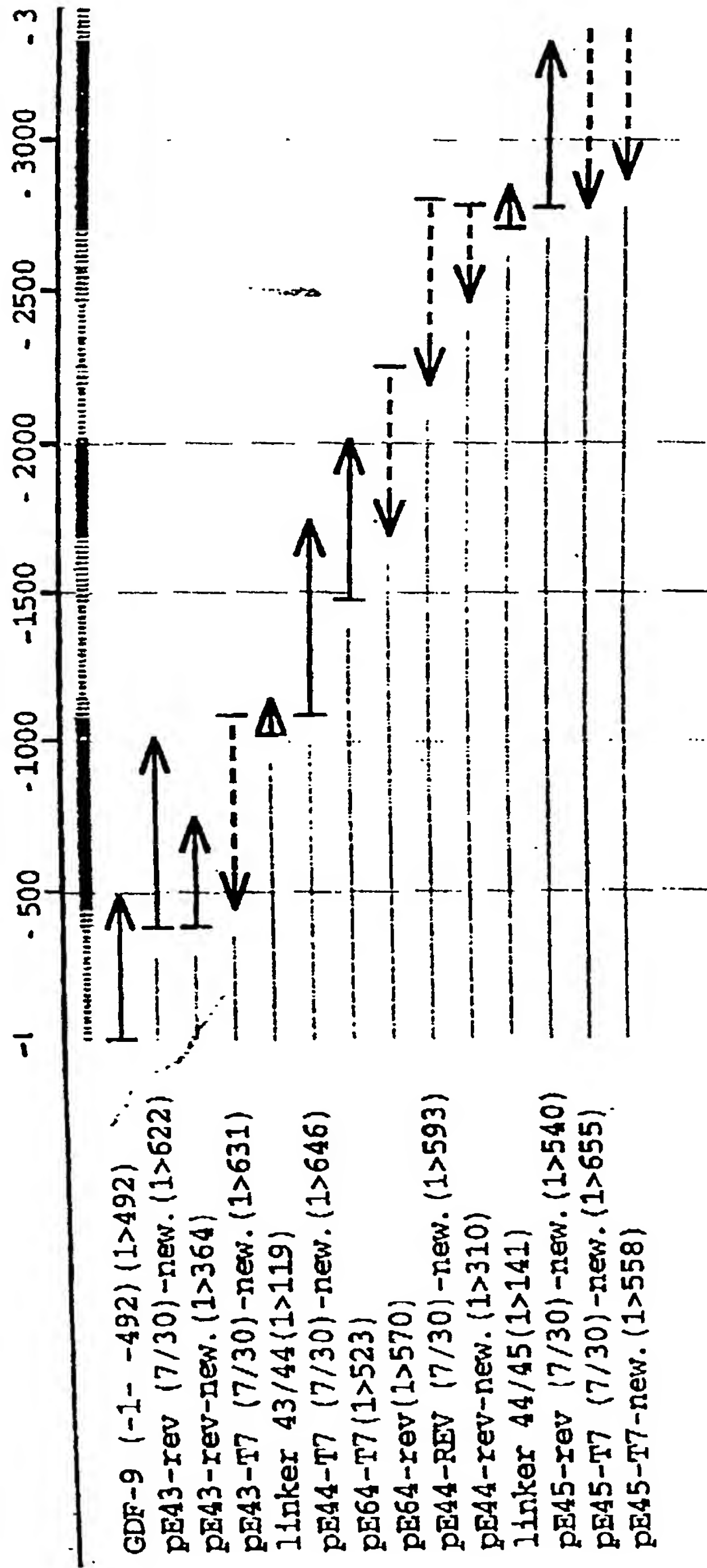


FIGURE 7

**DECLARATION, PETITION AND POWER OF ATTORNEY
FOR PATENT APPLICATION**

(Check one):

- ☐ Declaration Submitted with Initial Filing
☒ Declaration Submitted after Initial Filing

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

**GROWTH DIFFERENTIATION FACTOR-9 REGULATORY SEQUENCES
AND USES THEREFOR**

the specification of which (check one):

☐ is attached hereto.

OR

☒ was filed on **31 March 1999** as PCT International Application Number **PCT/US99/01785** and

☐ and was amended by PCT Article 19 Amendment on _____
(if applicable),

☐ and was amended by PCT Article 34 Amendment on _____
(if applicable).

I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

PRIORITY CLAIM

(Check one):

- ☐ no such applications have been filed.
- ☒ such applications have been filed as follows

1) FOREIGN PRIORITY CLAIM: I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (dd,mm,yyyy)	Priority Not Claimed	Certified Copy Attached	
				Yes	No
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

☐ Additional foreign application numbers are listed on a supplemental priority sheet attached hereto.

2) PROVISIONAL PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

Provisional Application Number(s)	Filing Date (dd/mm/yyyy)
60/080,108	01 April 1998 (01.04.98)

☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

3) U.S./PCT PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code, §120 of any United States application or §365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose information which is known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application Number	PCT Parent Number	Parent Filing Date (dd/mm/yyyy)	Parent Patent Number (if applicable)

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority sheet attached hereto.

POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorneys and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

James E. Cockfield	Reg. No. <u>19,162</u>	Jeremiah Lynch	Reg. No. <u>17,425</u>
Thomas V. Smurzynski	Reg. No. <u>24,798</u>	David J. Rikkers	Reg. No. <u>43,882</u>
Ralph A. Loren	Reg. No. <u>29,325</u>	Maria C. Laccotripe	Limited Recognition
Giulio A. DeConti, Jr.	Reg. No. <u>31,503</u>		Under 37 C.F.R. § 10.9(b)
Ann Lamport Hammitte	Reg. No. <u>34,858</u>	Debra J. Milasincic	Reg. No. <u>46,931</u>
Elizabeth A. Hanley	Reg. No. <u>33,505</u>	David R. Burns	Reg. No. <u>46,590</u>
Amy E. Mandragouras	Reg. No. <u>36,207</u>	Sean D. Detweiler	Reg. No. <u>42,482</u>
Anthony A. Laurentano	Reg. No. <u>38,220</u>	Cynthia L. Kanik	Reg. No. <u>37,320</u>
Kevin J. Canning	Reg. No. <u>35,470</u>	Theodore R. West	Reg. No. <u>47,202</u>
Jane E. Remillard	Reg. No. <u>38,872</u>	Shayne Y. Huff	Reg. No. <u>44,784</u>
DeAnn F. Smith	Reg. No. <u>36,683</u>	Hathaway P. Russell	Reg. No. <u>46,488</u>
Peter C. Lauro	Reg. No. <u>32,360</u>	Daniel B. Ko	Reg. No. <u>47,332</u>
Jeanne M. DiGiorgio	Reg. No. <u>41,710</u>	John S. Curran	Reg. No. <u>50,445</u>
Megan E. Williams	Reg. No. <u>43,270</u>		

Send Correspondence to:

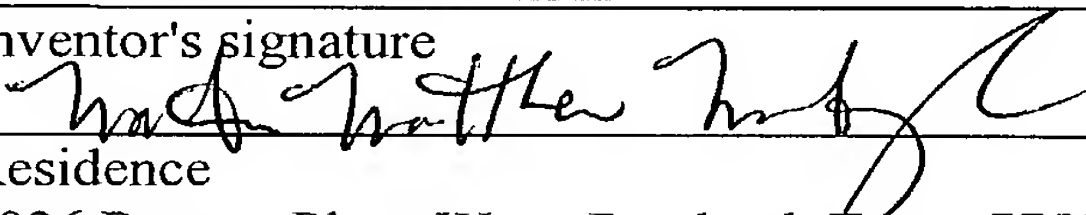
Jane E. Remillard, Lahive & Cockfield, LLP, 28 State Street, Boston, Massachusetts 02109, United States of America

Direct Telephone Calls to: (name and telephone number)


Jane E. Remillard, (617) 227-7400

Wherefore I petition that letters patent be granted to me for the invention or discovery described and claimed in the attached specification and claims, and hereby subscribe my name to said specification and claims and to the foregoing declaration, power of attorney, and this petition.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor <u>Martin Matthew MATZUK</u>	
Inventor's signature 	Date <u>5/16/02</u>
Residence 2926 Russett Place West, <u>Pearland</u> , Texas 77584 TX	
Citizenship United States of America	
Post Office Address (if different) Same as above	

2 - 00

Full name of second inventor Julia Andrea ELVIN	
Inventor's signature 	Date 5/21/02
Residence 3114 Stanton Street, <u>Houston</u> , Texas 77025 TX	
Citizenship United States of America	
Post Office Address (if different)	

- 1 -

SEQUENCE LISTING

<110> METAMORPHIX, INC. et al.

5 <120> GROWTH DIFFERENTIATION FACTOR-9 REGULATORY SEQUENCES AND USES THEREFOR

<130> MTN-029PC

<140>

10 <141>

<160> 5

<170> PatentIn Ver. 2.0

15

<210> 1

<211> 3379

<212> DNA

<213> Mus musculus

20

<400> 1

atgccttgga agaactaaga aggaacgcat ctctcgtgc cagtcttctt tctaaaggctc 60
ggggaggccc tgcggtggctc cttaaataaa atctaggtgt ggaagactga cttcacttct 120
gtaatcagac ctcaagcatg cctagctgta cattgtttct tatcagctct acagcagtca 180
25 gctgacaaca cggttttatc caatttgctc tcattagata cagatttctc tacttatttg 240
gctttcttct cctgtttctc ctggacattt atttcaaacc caccaagtat cgtcacccta 300
gtggactact tatataggtg agagcccaac agaaacccaa ggaattataa aacaaaacct 360
cttagggctg gagatgtgac tgcagaagta ttaggtttat ttttctctct tgattcatcg 420
tctatctcca cagctcccac gtccaggtgg ttagaataac tgatggctta ctgcacagtg 480
30 aatcagagaa acaaagtcac aatgggttct gaaattccag tgatgttagg ctgttcagcc 540
tgtccaggtg gcttagtaga ggggttgtaa agttggccca agacaacagc tacttgcaat 600
gtgctttgaa aaaggacatc agggcccagg ttagagaggtt cagccttta atcccagcat 660
tcaggagaca caggcaggtg gatctctgtg gattcaagcc agcctggtct atatagagag 720
ttctaggaca gtcagagcta cataattttg agagacccca tctcaaaata aaaagaggac 780
35 atcaccttta actttgagca attaaaacat acaatgtctt ggggccagaa agatgatgga 840
tgacctggca gttaagggca ctggttgctc ttcttagaac aaccatttat agctccactt 900
ccagggggag ccaatgcctt tctcgggcat ccttaggaat acaaaaacac ccattataaa 960
atgaaaataa acaattaaaa aaaagccaca atatctagct ggaatggcaa gatagcaact 1020
tgccgccaag cctaacgacc aaaagtcac cagtcctaaa ggatcaagct tattaagctt 1080
40 gatcctaact tcaaccctg tgatccacag agaaaaaggg agaagagact tetacagctc 1140
gtcttctgac ctacaccctc agcacttaca taccaaatac atattttaaa atgtagtgac 1200

WO 99/50406

PCT/US99/07185

- 2 -

agtacaatat cctgatgccc ttgttttggg ctacaattca cttgactata ctttggttga 1260
 aaaatcttga gacatcagca aatgcaatta aatagacatc gtattttattc ttttaaccaac 1320
 caccagatg gcagaaatca ggacagtgtt ttagtccttg attaaagtat ttgtgtcccc 1380
 gccttggagg tgagatcaag ctagataatg ttttaaggaaa atggggggcag ggtggaagtg 1440
 5 gggggggagaa tgtgcagtgc accatgaaag gctatctagc accccttgga aggagactat 1500
 ctcttgggag aaagtcaaaa cacactaatg acacaggata aaaaaaaaaat cagtttgccc 1560
 agcaccagat agataaggct gaatctcttg caaaactgaa ctacccgcca aactgccccca 1620
 gtttgaattg ccaaattatt tacagcccag ttttaaccctt tacagccctg agaggaggag 1680
 aaagggcggt gagtcccaca ggataatatt agctaggaga atgaatgcac ttttagagaaa 1740
 10 acataacagt tccttagaga gtgggggtctg ggtttgatga cacactcagc ccaaaccacac 1800
 cactttctcc ttggctcccc tcaccattat gtgcgtctga caaactgcca catcgttcgt 1860
 ggctgaactg ggagaaacgt ttatgtcaca tcttggaaca gtctacagcc tgccatggct 1920
 tgtaggcagt gtgtacccaa gctctagatc cgtgtcctta attttgtact gtggtcagtg 1980
 aaggcccttt tggttatgag gaccacatat acctgggtgga gagtaaggga cagcagaaga 2040
 15 acgaggagaa caagtctttt tgttttttct ttttaatttt tgagcagctg ccggtctgta 2100
 tgcaaaggcc tggtagtgag cggaaaggac cgcctcttgg aacacccagg tgctcctaag 2160
 cagcccacac acctaatecg cgcctcatga cctataggcc ccgcccctg cacttctcga 2220
 cgggggacct tacggacgta tccacctaa atccccaggg tggctaataa gccccgcca 2280
 cttgcttggc acccttcccc ccaagccacc attcctcttc aagccccgcc ccttggcaac 2340
 20 atcaacctct ccggtgcgtt cactacttac cgtaagccca caagcacctc tataagcccc 2400
 acccacgtgt cgctagaagc gcctctaagc tccgcccagc ctcttggtat ggtcctttcc 2460
 ggctacacga aggtcaggcc acgcccactc aggtcaactg acaatttttt ttttgagtct 2520
 ccccttgag atcgaaagaa aatggtagtt ccggaatg acttcacgtt ctctaggctc 2580
 ggcttccgct cctcccggaa gtgacgggag aaagccggct cagagctctg gaggtagagg 2640
 25 tgaccgatcc tgggggtcgg tgttaggccg tgggagggtt ttcgcgagac tgagccacgc 2700
 gtctatcttc tgtcccaggg ccgcccagga attcgacatg ggccgcgagt ttgggaacct 2760
 ggccgagata cggcacgtga tctcctacag cttgtgcgcc tttgagcagc gcgccttccc 2820
 aagctatttc agcaaaggca tccccaacgt gtgcgcgcga ctgcgcgagc catcctgcgc 2880
 gtggcgccgc gtgagtgcgg ggccggcggg gtacagccgc gcacctctcc ttcgcagcc 2940
 30 ctgcgggggc agacgtggag gagttcttat catatcgaaa tgactggcct aaggtcactt 3000
 gggtagctcc aacggctccc tctgagggtca tcgctggcca ctatctgtga tgaaaggctg 3060
 atgcaatcag tgtcttgtgt tagatgtgcg gccgattttg ctgtcgggtt catggcattg 3120
 tgaatatccg tgaagggtgca cttaaatttg atggagaggt tgaggcccga tcttgctgta 3180
 gaagccaggc tggccttaaa cttgagattt tactgcctc tectcgggac tgttggtgg 3240
 35 aattgtaggc atgccgccag cacaacagta ctttttaaaa tggtagaat tatgttaggt 3300
 aaattccgtg attttttttt ttttaatttg gtgttaacta ggaaagcgaa ggccacatct 3360
 tctgggctg atatcaagc 3379

<210> 2

40 <211> 370

<212> DNA

<213> Mus musculus

<400> 2

cttgggaagaa ctaagaagga acgcacatcc tgctgccagt cttcttttcta aaggctcgggg 60
 5 aggcctctgcg gtggctcctta aataaaatct aggtgtggaa gactgacttc acttctgtaa 120
 tcagacctca agcatgccta gctgtacatt gtttcttata agctctacag cagtcagctg 180
 acaacacggg tttatccaat ttgtttctcat tagatacaga tttatctact tatttggtctt 240
 tctttccctg tttctcctgg acatttattt caaaccacc aagtatcgtc accctagtgg 300
 actacttata taggtgagag cccaacagaa acccaaggaa ttataaaaca aaacctctta 360
 10 gggctggaga 370

<210> 3

<211> 362

<212> DNA

15 <213> Homo sapiens

<400> 3

cttgggagaa ctagtgagga acatatttct ccatgccagt cctcttctta aaggccagga 60
 agagcctagc ttgggtctctt aaataaattt caggtgtgtg ggtggactac ggtcacttct 120
 20 gtaatcaggc ctcaagtatg cctagctgaa gattttatca gctctatata aagcagctga 180
 taacacctta tttagccaat ttgttaatta gatacagatt tacttggtta tttagcttct 240
 ctctcttttc ccttggcatt tacgtaaaac ccgttactgt tacactaccg gactaactat 300
 gtagctgaaa gtctaagata aattgaggaa tgatcttaaa taaaatctct cattttttga 360
 ga 362

25

<210> 4

<211> 26

<212> DNA

<213> synthetic primer

30

<400> 4

ttcttcaagg gatccactt tccag 26

<210> 5

35 <211> 25

<212> DNA

<213> synthetic primer

<400> 5

40 ctgggaagtg ggatccctgg aagaa 25